SchizoREmics

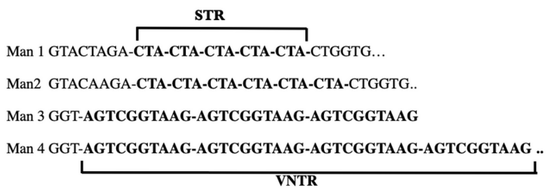
Rachele R. Rubiu

Introduction

Schizophrenia is a severe long-term mental health condition, characterized by continuous episodes of psychosis. Major symptoms include hallucinations (typically hearing voices), delusions, and disorganized thinking.

Risk factors
- Genetics
- Environment

→ The repeat variants



Hypotesis

1. Repeat variants in genes involved in brain functions play a role on the risk for schizophrenia
2. The repeat variants with CG motifs have an effect on the DNA methylation
3. The aberrant DNA methylation in CG repeat variants impacts the gene transcription by changing their expression levels or by leading to allelic imbalance
4. The repeat variants impact ultimately the clinical phenotype (i.e. the severity of the symptoms or the cognitive deficits) and the cellular phenotype

Experimental approach

Identification of the repeat variants

1. Calling of the repeat variants.

Expansion hunter ExpansionHunter Denovo HipSTR CNVnator

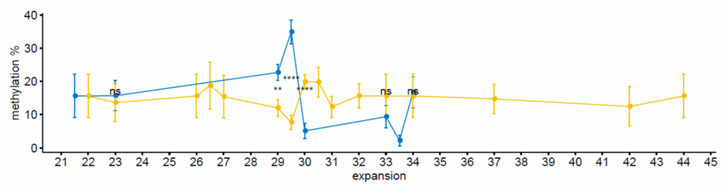
2. Annotate them based on:
- their prevalence in the general population
- their localization in specific regions.
- their nucleotide content.
- their tendency to impact the gene expression, to modify the DNA methylation

3. Identify the outliers

Results
A list of the candidate repeat variants as well as a list of patients carrying them.

Determination of the impact of the candidate repeat variants on DNA methylation

1. Correlation of candidate CG repeat variants with DNA methylation
We will be using BS Seeker2 and CGmapTools to align and generate CGmap from the RRBS data, to subsequently generate a matrix of methylation.Then we will correlate the DNA methylation levelwith the size of the CG repea variants in each locus of interest



2. Identify the longitudinal changes of DNA methylation

3. Identification of epivariations
An Independent F-tests will be conducted in each CpG.
A list of potentially pathogenic epivariations will be selected if:

- the DNA methylation level display greater variance after the follow-up compared to baseline (variance increases with time),
- the CpG is identified only in the affected group (CpG more variable in psychosis), and
- DNA methylation level deviated from more than 10% between the two time points

Results
Discovering the percentage of stability of the DNA methylation in candidate repeat variants and having a list of individuals carrying epivariation

Determination of the impact of the candidate repeat variants on gene expression

1.Identification of the impact of repeat variants on the level of gene expression
Correlate the level of expression of the genes near the repeat regions with the size of the repeat

2. Identification of the impact of repeat variants on the allelic specific expression
Allele-specific expression analysis from the RNAseq data by calculating read counts per allele.

Results
List of repeat variants with a functional effect on gene expression and a list of individuals with altered gene expression

Influence on the clinical phenotype

1. Over-representation of clinical groups among the 3 lists of outliers
The overlap between the lists of repeat-individual match, methylation-individual match, and the expression-individual match will be assessed using a Venn diagram.

2. Correlation of the pathogenicity of the candidate repeats with clinical dimensions.

Results
Identification of possibly-pathogenic repeat variants, estimation of the contribution of repeat variants to the risk for schizophrenia and release of a pipeline to detect possibly-pathogenic repeat variants

Impact of repeat expansions in the brain

1.Testing the relationship between expansions, DNA methylation and gene transcription in the brain
Using BECon to assess the blood-brain correlation by using DNA methylation data from paired samples between the whole blood and three human brain regions. For our best candidate expansions with CG motifs, we will use this tool to estimate the correlation between the DNA methylation level detected in our blood samples and the brain.

	Chr	Coor	Gene(s)	Gene Region(s)	Correlation		
cg07147350	23	146992908	FMR1-AS1, FMR1	intragenic, promoter	0.81	0.48	0.59
cg22625568	23	146993010	FMR1-AS1, FMR1	intragenic, promoter	0.55	0.57	0.56
cg05288927	23	146993092	FMR1-AS1, FMR1	intragenic, promoter	0.62	0.49	0.7
cg14470409	23	146993125	FMR1-AS1, FMR1	intragenic, promoter	0.33	0.54	0.4

Result
Estimation of the brain-blood correlation of the expansion-methylation relationship

Expected results

- to understand the structure and the effect of repeats sequence in human genome, by identifying hot spots for repeat variants in genes playing a role in brain function or psychiatric disorders
- to increase the yield of molecular diagnosis in schizophrenia
- to identify factors influencing the risk to develop schizophrenia as well as the severity of the phenotype
- to better understand how aberrant DNA methylation or abnormal RNA transcription in adolescence leads to the onset of psychosis
- to identify new therapeutic targets for intervention in patients suffering from psychosis
- to better understand the cellular impact of repeat variants